第28卷 第1期
2011年02月

工 程 数 学 学 报
CHINESE JOURNAL OF ENGINEERING MATHEMATICS

Vol. 28 No. 1
Feb. 2011

文章编号:1005-3085(2011)01-0015-06

一种改进的多通道盲信号解相关算法*

徐仔波[1]， 陈文超[1]， 高静怀[1]， 邸双亮[2]

(1- 西安交通大学电子与信息工程学院，西安 710049； 2- 西安交通大学理学院，西安 710049)

摘 要: 针对多通道解相关算法在混合信号协方差矩阵的最大特征值较大或者最大与最小特征值的比值很大 (大于 10^6 的病态混合) 时收敛速度变慢且收敛误差增大的不足，本文引入归一化方法，即在迭代过程中对解混信号进行归一化，从而限制其协方差矩阵最大特征值的取值范围，且降低最大与最小特征值的比值. 数值仿真表明，改进后的算法在降低迭代误差和加强稳定性方面有了明显的改善，收敛速度也得到了可观的提升，改进算法也具有更广的适应范围.

关键词: 解相关；特征值；病态混合；归一化；盲信号分离；独立成分分析

分类号: **AMS**(2000) 94A99　　**中图分类号**: TN911.73　　**文献标识码**: A

1 引言

近年来，盲信号分离在通信信号处理、生物医学信号处理、语音信号处理等领域得到了广泛的应用[1-3]. 一些常用于盲分离的自适应算法，例如，独立成分分析，要求对混合信号进行预白化 (球化). 一个零均值随机矢量称为白的，是指其协方差阵是一个单位矩阵. 在白化时，信号矢量 $x(k)$ 用下列变换进行预处理

$$y(k) = Wx(k), \tag{1}$$

目的是选择矩阵 W，使得信号 $y = Wx$ 的协方差矩阵 $E\{y(k)y(k)^T\}$ 变成单位阵[4]. 白化预处理在自适应滤波以及多层神经网络等系统中也常被用作预处理过程来提高自适应性能，因为它能简化计算复杂度，在独立成分分析中，它能减少近一半未知参数的计算[2].

通常情况下，多通道解相关算法的预白化过程稳定且快速高效[5-7]. 然而，当混合信号协方差矩阵的最大特征值较大以及最大与最小特征值比值很大 (大于 10^6 的病态混合) 时，会出现收敛性能恶化的现象. 针对这一问题，我们在多通道解相关算法中引入了对每一步解混信号的归一化，从而控制了最大特征值的取值范围，降低了最大与最小特征值的比值，这样改进后的算法在降低迭代误差和加强稳定性方面有很大的提高，同时收敛速度提高了至少一个数量级.

本文第 2 部分简要介绍了全局和局部自适应两种多通道解相关算法；第 3、4 部分分析了算法的性能，并针对原算法的不足提出了一种改进方法，另外还给出了改进算法的步长系数的经验公式；最后一部分是改进前后算法的数值仿真对比.

收稿日期: 2008-04-10. **作者简介**: 徐仔波 (1984年7月生)，男，硕士. 研究方向：盲信号处理及其在地震信号处理中的应用.

***基金项目**: 国家自然科学基金 (40674064; 40730424)；国家高技术研究发展计划 (2006AA09A102-11)；国家科技重大专项 (2008ZX05025-001-009; 2008ZX05023-005-008).

2 独立成分分析与多通道解相关算法

独立成分分析在盲信号分离中占据着重要的地位，其混合模型为 $x = As + v$，其中 x 为混合信号，A 为未知混合矩阵，$s = [s_1, \cdots, s_n]^T$ 为相互独立的源信号，v 为观测噪声。其分离模型为 $y = Wx$，其中 W 为分离矩阵. 一般还假设 s 的各个分量的均值为零、方差为1且 A 为列满秩矩阵等. 它利用一个神经网络 W 来对混合信号 x 进行处理，通过无监督的自适应学习来优化一个对照函数 (Contrast function)，并不断调整网络权值 W，使得 W 和混合矩阵 A 的乘积逐渐收敛至一个广义排列矩阵 H，此时网络输出 y 即为源信号波形的估计值[8]. 在上述过程中，白化预处理过程显得尤为重要，而常用的预白化处理算法就是多通道解相关算法.

典型的多通道解相关算法有 Silva 和 Almeida[5,6] 所提出的全局自适应算法

$$W(k+1) = W(k) + \eta(k)\big(I - y(k)y(k)^T\big)W(k), \tag{2}$$

以及 Cichocki, Kasprzak 和 Amari 提出的局部自适应算法[8]

$$W(k+1) = W(k) + \eta(k)\big[I - y(k)y(k)^T\big]. \tag{3}$$

两种算法都涉及一个问题：如何选取 $W(0)$ 和 $\eta(k)$，从而使算法有最佳的收敛性能？Douglas 和 Cichocki 给出了上述两种算法的误差演变式以及步长系数的选取方法，使得系数矩阵 $W(k)$ 能够快速、稳定、稳健地收敛[9].

对于全局自适应算法[9] 有

$$\tilde{\lambda}_i(k+1) = \big[\big(1 - \eta(k)\lambda_i(k)\big)^2 - \eta(k)^2\lambda_i(k)\big]\tilde{\lambda}_i(k). \tag{4}$$

上式是对 (2) 式的误差演变分析，其中 $\tilde{\lambda}_i(k) = \lambda_i(k) - 1$，$\lambda_i(k)$ 为 $\bar{R}_{yy}(k)$ 的第 i 个特征值，$\bar{R}_{yy}(k)$ 的定义见文献 [9] 中式 (19)，$\eta(k)$ 为步长系数. 由 (4) 式，Douglas 和 Cichocki 得出了保证每个 $\lambda_i(k)$ 稳定地收敛到1的、$\eta(k)$ 应满足的充分条件

$$0 < \eta(k) < \eta_{\max}(k), \tag{5}$$

$$\eta_{\max}(k) = \begin{cases} \frac{1}{\sqrt{\lambda_{\max}(k)}}, & 若\ 0 < \lambda_{\max}(k) \le 3 + 2\sqrt{2}, \\ \frac{2}{\lambda_{\max}(k) - 1}, & 若\ \lambda_{\max}(k) > 3 + 2\sqrt{2}, \end{cases} \tag{6}$$

此即文献 [9] 中的 (22) 与 (23) 两式，式中 $\eta_{\max}(k)$ 为最大步长系数，$\lambda_{\max}(k)$ 为 $\bar{R}_{yy}(k)$ 的最大特征值.

对于局部自适应算法 (3)，见文献 [9] 的式 (28)，有

$$\tilde{\lambda}_{w,i}(k+1) = \big[1 - \eta(k)\sqrt{\lambda_i}\big(1 + \sqrt{\lambda_i(k)}\,\big)\big]\tilde{\lambda}_{w,i}(k), \tag{7}$$

上式是对算法 (3) 的误差演变分析，其中 $\tilde{\lambda}_{w,i}(k) = \lambda_{w,i}(k) - 1/\sqrt{\lambda_i}$ 定义为 $W(k)$ 的误差，$\lambda_{w,i}(k)$ 为 $E\{W(k)\}$ 的第 i 个特征值，λ_i 为 R_{xx} 的第 i 个特征值，$\lambda_i(k) = \lambda_i\lambda_{w,i}^2(k)$. 利用式 (7)，Douglas 和 Cichocki 得出了保证 $\lambda_{w,i}(k)$ 收敛到 $1/\sqrt{\lambda_i}$ 的、$\eta(k)$ 应满足的稳定性条件

$$0 < \eta(k) < \frac{2}{\sqrt{\lambda_{\max}}\big(1 + \sqrt{\lambda_{\max}(k)}\,\big)}, \tag{8}$$

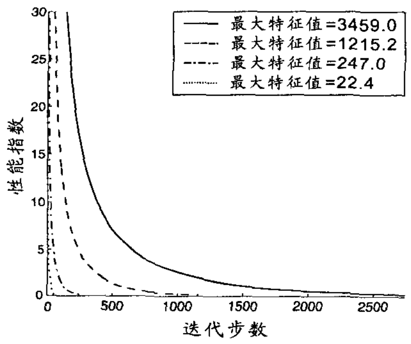
其中 $\lambda_{\max}$ 和 $\lambda_{\max}(k)$ 分别为 R_{xx} 和 $\bar{R}_{yy}(k)$ 的最大特征值.

3　算法性能分析

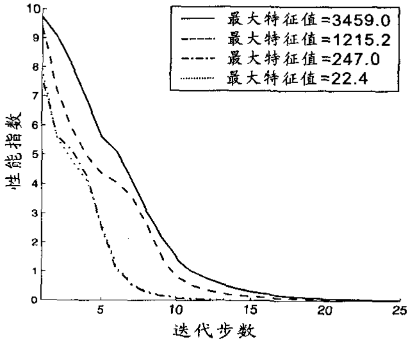
当上述$\lambda_{\max}$和$\lambda_{\max}(k)$很大时，由式(6)和式(8)可知，步长系数$\eta(k)$取值变得很小，此时收敛速度变慢，如图1(a)、图2(a)所示．特别地，当混合信号协方差矩阵的最大特征值$\lambda_{\max}$与最小特征值$\lambda_{\min}$之比(定义$\Delta = \lambda_{\max}/\lambda_{\min}$)很大(大于$10^6$的病态混合)时，若初值$W(0)=I$，假定$\bar{R}_{yy}(k)$初始变化不大，即$\bar{R}_{yy}(k)\simeq R_{xx}$，所以误差演化式(4)和式(7)可以近似为

$$\begin{cases} \tilde{\lambda}_i(k+1) \simeq \left[\left\{1-\frac{2\lambda_i(k)}{\lambda_{\max}}\right\}^2 - \frac{4\lambda_i(k)}{\lambda_{\max}^2}\right]\tilde{\lambda}_i(k) = \alpha\tilde{\lambda}_i(k), \\ \tilde{\lambda}_{w,i}(k+1) \simeq \left[1-\frac{2\sqrt{\lambda_i}(1+\sqrt{\lambda_i(k)})}{\lambda_{\max}}\right]\tilde{\lambda}_{w,i}(k) = \beta\tilde{\lambda}_{w,i}(k), \end{cases} \tag{9}$$

显然，对于那些小特征值$\lambda_i \simeq \lambda_{\min}$的误差传替项$\alpha$和$\beta$($\alpha$和$\beta$由上式确定)将接近1．此时，$\bar{R}_{yy}(k)$将停留在距离单位矩阵$I$的一定误差处而不收敛至$I$，如图3所示．

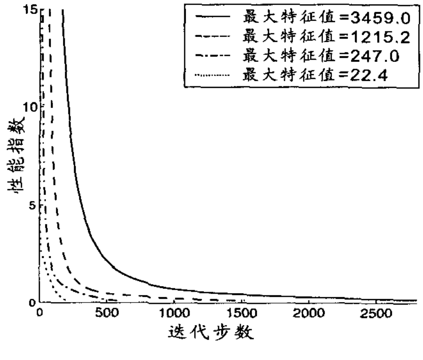


(a) 改进前收敛图

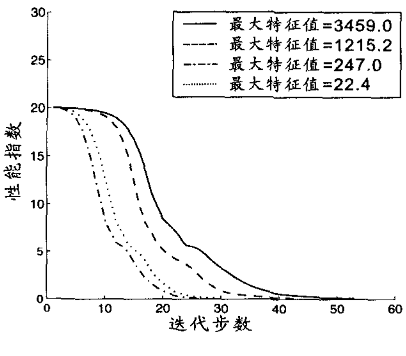


(b) 改进后收敛图

图 1:　改进前后的全局自适应算法应用于不同最大特征值的混合信号收敛图

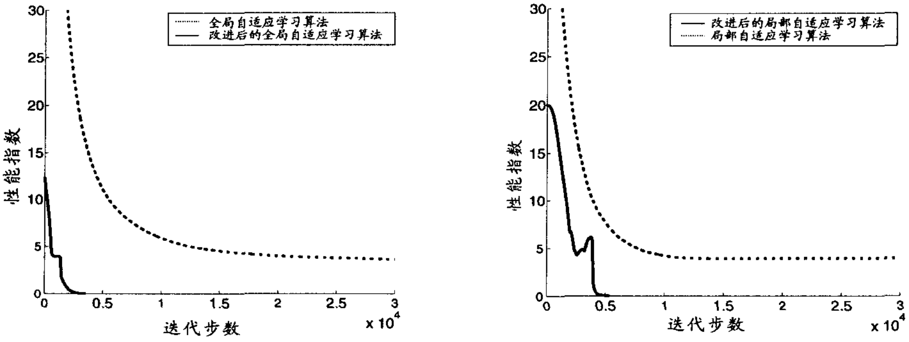


(a) 改进前收敛图



(b) 改进后收敛图

图 2:　改进前后的局部自适应算法应用于不同最大特征值的混合信号收敛图



(a) 全局自适应学习算法　　(b) 局部自适应学习算法

图 3: 改进前后全局和局部自适应算法应用于 $\Delta_x = 1.022 \times 10^7$, $\lambda_{\max} = 35076$ 的混合信号收敛图

4 多通道解相关改进算法

我们对于现有的解相关算法提出如下改进：在每一步解相关迭代之前，将解混信号 y 进行归一化

$$z(k) = G(k)y(k), \tag{10}$$

其中

$$G = \operatorname{diag}\Big(\Big\{\sum_{t=1}^{T}\big[y(1,i)\big]^2\Big\}^{-\frac{1}{2}},\cdots,\Big\{\sum_{t=1}^{T}\big[y(n,i)\big]^2\Big\}^{-\frac{1}{2}}\Big), \tag{11}$$

即解混信号 y 在时域上的归一化矩阵，$\operatorname{diag}(\cdot)$ 表示取 $(\cdot)$ 中元素为对角元素的对角矩阵，T 为信号时域长度，n 为信号维数. 显然，归一化对于 y 中的大功率 (大特征值) 分量有压制作用，而对于小功率 (小特征值) 分量有放大效果，即归一化前后 $\Delta_z(k) \le \Delta_y(k)$，加快了收敛速度. 此时，$z$ 的每一行元素的平方和为 1，$\bar{R}_{zz}(k) = E\{z(k)z(k)^T\} \simeq z(k)z(k)^T$ 是一个对称正定矩阵，主对角元素为 1，$\bar{R}_{zz}(i,i) = 1 > |\bar{R}_{zz}(i,j)|$，其中 $i,j = 1,\cdots,n$ 且 $i \ne j$. 所以，对于具有大的 $\lambda_{\max}(k)$ 的 $\bar{R}_{yy}(k)$，上述变换将最大特征值降到一个合适的水平. 可以证明，当 z 的维数为 n 时，$\bar{R}_{zz}(k)$ 的最大特征值 $\lambda_{z,\max}(k)$ 满足 $0 < \lambda_{z,\max}(k) < n$. 这可以解决收敛速度随最大特征值增大而减慢的问题，且由于最大特征值 $\lambda_{z,\max}(k) \in (0,n)$，根据式 (6)、式 (8)，步长系数 $\eta(k)$ 的取值范围 $(0,+\infty)$ 将变成 $(+\xi,+\infty)$，而 ξ 的取值跟信号维数有关. 于是，将 z 代入上述自适应学习算法得到如下改进算法

$$W(k+1) = W(k) + \eta(k)\big[I - G(k)y(k)y(k)^TG(k)\big]W(k), \tag{12}$$

$$W(k+1) = W(k) + \eta(k)\big[I - G(k)y(k)y(k)^TG(k)\big], \tag{13}$$

其中 $y(k) = W(k)x(k)$，G 如式 (11) 所示.

通过大量的数值仿真实验，我们给出改进后的全局和局部自适应算法的最佳步长系数 η_{opt} 的经验公式

$$\eta_{opt} = \begin{cases} (n+1)^2\big(n\sqrt{\Delta_x}\big)^{-1}, & \text{全局算法}, \\ (n+1)^2\big(2n\sqrt{\Delta_x}\big)^{-1}, & \text{局部算法}, \end{cases} \tag{14}$$

其中 n 为信号维数，Δ_x 为协方差矩阵 R_{xx} 的最大特征值与最小特征值之比.

5 数值仿真

1) 病态混合以及不同最大特征值的混合信号的解相关过程

在Matlab环境下，选取源信号 $s = [s_1, s_2, s_3, s_4, s_5]^T$ 且 $s_i^2 = 1$，取混合矩阵 A 为随机矩阵进行混合的同时利用调幅参数 a_i 改变各源信号幅值，从而得到不同最大特征值 $\lambda_{\max}$ 和大 Δ_x 值的混合信号 x (病态混合)，即 $x = A[a_1s_1, a_2s_2, a_3s_3, a_4s_4, a_5s_5]^T$.

图1、图2中四条收敛曲线对应混合信号 x 的最大特征值 $\lambda_{\max}$ 依次为 3459.0, 1215.2, 247.0, 22.4, Δ_x 依次为 2169, 1710, 598, 559. 根据式(5),(8)以及式(14)，图1(a)、图1(b)和图2(a)、图2(b)中最大特征值 $\lambda_{\max}$ 从大到小对应学习步长系数 $\eta(k)$ 依次取为 0.00056, 0.0016, 0.0081, 0.061；0.105, 0.11, 0.22, 0.19；0.00028, 0.0008, 0.0042, 0.05；0.06, 0.07, 0.12, 0.10. 由图1、图2可知，改进算法步长系数 $\eta(k)$ 并不随最大特征值 $\lambda_{\max}$ 变化而急剧变化，而是在小范围内波动，收敛稳定且快速. 全局自适应算法的收敛步数随最大特征值在100-2500步内变化，最大特征值越大收敛越慢；改进后全局自适应算法的收敛步数大都在20步左右. 局部自适应算法的收敛步数随最大特征值在250-2500步内变化，最大特征值越大收敛越慢；改进后局部自适应算法的收敛步数大都在30-50步左右. 图3显示了改进后的算法在最大特征值与最小特征值的比值较大情况下的收敛性能.

根据式(5),(8)和式(14)估计最佳学习步长系数，图3(a)、图3(b)中改进前后的全局和局部自适应算法学习步长系数依次取为 0.000056, 0.0011 和 0.000025, 0.0007. 显然，改进后的算法不会出现收敛误差大的问题且收敛速度加快. 全局算法的收敛步数达到30000时，依然存在较大误差，且很难继续收敛至单位矩阵，而改进后全局算法大约在3500步左右收敛至理想结果；如图3(b)所示，局部算法与全局算法类似，改进的局部算法大约在5000步左右收敛至理想结果.

2) 运算时间的分析与统计

局部自适应算法：原算法中存在2次矩阵相乘；在改进算法中，存在3次矩阵相乘，这里 $G(k)$ 的时间计算复杂度可以忽略不计. 全局自适应学习算法：原算法中存在3次矩阵相乘，而改进算法中存在4次矩阵相乘. 因此，对于局部自适应算法，改进后每一步的时间计算复杂度大概提高了50%，而全局算法大概仅提高了33%，然而，迭代步数的减少量在 $10 \sim 10^3$ 量级，单步的时间复杂度对于整个算法的运算速度的影响就可以忽略不计了。

计算环境如下：计算机型号为HP Compaq dc 7800 Convertible Minitower；处理器型号为Intel(R) Core(TM)2 Quad CPU Q6600 @2.40GHz；内存(RAM)为4.00GB，得到了如表1所示的统计结果. 虽然在单步运算时间上改进算法更长，但整体耗时要比原算法少. 实际运算时间与上述的分析基本一致(表中全局算法的平均每步耗时稍小于局部算法的平均每步耗时，可能与计算机的稳定性有关). 特别地，对于某些大特征值混合信号的解相关过程，改进算法运算时间远远比原算法少，例如对于 $\lambda_{\max}$ 为3459.0时，如表1所示.

表1: 不同解相关算法的运算时间统计

特征值	3459.00	1215.20	247.00	22.40	平均步时
局部算法	2745步/0.24 s	1532步/0.17 s	554步/0.063 s	52步/0.0061 s	0.000098 s
全局算法	2633步/0.22 s	1182步/0.12 s	248步/0.029 s	56步/0.0067 s	0.000092 s
改进局部算法	53步/0.014 s	43步/0.011 s	27步/0.0073 s	28步/0.0075 s	0.00027 s
改进全局算法	22步/0.0060 s	22步/0.0060 s	14步/0.0040 s	14步/0.0038 s	0.00027 s

6 结论

本文在全局和局部两种解相关算法中引入了对解混信号的归一化，限制了解混信号协方差矩阵的特征值(控制其协方差矩阵的最大特征值取值范围且降低了最大与最小特征值之比值)，使算法在降低误差和加快收敛速度方面得到了提高．同时，改进后的算法具有相对稳定的收敛性能，能适应于更广的应用范围．

参考文献：

[1] 乔永凤，马建芬. 独立分量分析技术在语音信号处理中的应用[J]. 电脑开发与应用，2007, 20(4): 22-23
Qiao Y F, Ma J F. Application of independent component analysis technology in speech signal processing[J]. Computer Development & Applications, 2007, 20(4): 22-23

[2] Hyvärinen A, Oja E. Independent component analysis: algorithms and applications[J]. Neural Networks, 2000, 13(4-5): 411-430

[3] Vigario R. Independent component analysis for identification of artifacts in magnetoencephalographic recordings[J]. Advances in Neural Information Processing Systems, 1998, 10: 229-235

[4] Cichocki A, Amari S I. Adaptive Blind Signal and Image Processing[M]. LTD: John Wiley & Sons, 2002

[5] Silva F M, Almeida L B. A distributed solution for data orthonormalization[C]// International Conference of Artificial Neural Networks, Espoo, Finland, 1991, 1: 943-948

[6] Silva F M, Almeida L B. A distributed decorrelation algorithm[J]. Neural Networks, 1991: 145-163

[7] Cichocki A, Kasprzak W, Amari S I. Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals[C]// International Symposium of Nonlinear Theory Application, Las Vegas, NV, 1995: 61-65

[8] Comon P. Independent component analysis-a new concept?[J]. Signal Processing, 1994, 36: 287-314

[9] Douglas S C, Cichocki A. Neural networks for blind decorrelation of signals[J]. IEEE Transactions on Signal Processing, 1997, 45(10): 2829-2842

An Improved Algorithm for Blind Multi-channel Signal Decorrelation

XU Zai-bo[1], CHEN Wen-chao[1], GAO Jing-huai[1], DI Shuang-liang[2]

(1- School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049;
2- School of Science, Xi'an Jiaotong University, Xi'an 710049)

Abstract: As the maximal eigenvalue of the mixture's covariance matrix becomes very large or the ratio of the maximal eigenvalue to its minimal eigenvalue is very large (which is larger than 10^6 as ill-conditioned mixture), the current blind multi-channel signal decorrelation algorithms tend to be slow in convergence and large in decorrelation error. In this paper, we propose a new normalization algorithm which restricts the range of the covariance matrix's maximal eigenvalue and depresses the ratio of the maximal eigenvalue to its minimal eigenvalue. The new algorithm can greatly improve the convergence speed, decorrelation error, and computational stability of the blind multi-channel signal decorrelation. The simulations verify these advantages of the proposed algorithm.

Keywords: decorrelation; eigenvalue; ill-conditioned mixture; normalization; blind signal separation; independent component analysis

Received: 10 Apr 2008. **Accepted:** 21 Sep 2010.

Foundation item: The National Natural Science Fundation of China (40674064; 40730424); the National High Technology Research and Development Program of China (2006AA09A102-11); the National Science and Technology Major Project(2008ZX05025-001-009; 2008ZX05023-005-008).